



The Dataflow chart for the NSR and SSR

Explanation Note

1	Process	Explanation	Responsibility
1	Household Data Collection	As coordinated by the CBTT leader, the enumerators collect household and household member information using the NASSCO checklist.	
2	Data upload from the tablets	Completed household's forms on CAPI application are uploaded from the tablets to the server.	
3	CAPI data server	This hosts the households data collected and stores it for retrieval. The on-the-shelf CAPI application data storage capacity is usually small and it either requires frequent data download and purging or connection to another secondary server.	
4	Crude Data Cleaning	<p>This will require the MIS or data officer at the state to do the following:</p> <ul style="list-style-type: none"> -Check for data field omission, field mismatch, improper or incorrectly entered responses. -Perform merging of individual levels data with households data -Assess completion time of survey and provide feedback as either – commendation, redo, or flag for a strike. - checking for duplicated reference number, household, household member id etc. - checking for duplicated records: whenever there is a need to redo an interview, there will be duplicated records for the same households. Therefore, this need to be taken care off and incomplete enumerations need to be deleted when the complete enumeration has been achieved. <p>-Instructions to enumeration could be:</p> <ol style="list-style-type: none"> 1. The following ___households in the _____ communities need redo 2. The following information is missing on these hhs ----- 3. I'm not sure this person in household # is the household head, verify. 	
5	Redo Household Data Collection	<p>As required, enumerators might have to:</p> <ol style="list-style-type: none"> 1. Redo the survey using CAPI and upload to the server 	

			2. Using a tracking sheet, enumerator will collect necessary few missing information.	
6	Crude Data Validation		This is provide a second layer of data quality checks to the states. The state MIS will forward the a round ¹ of data collected and corrected to the extent of their knowledge to the NASSCO team for a second layer data quality review. The response time from the NASSCO should not be more than 5 days.	
7	Backend Data Collection.	Data	This is a data validation exercise that seeks to pick fraud and inconsistence in household's enumeration. This is usually done by developing a small version of main questionnaire will stable questions from the original checklist that may not be more than 20 questions. This backcheck questionnaire is programmed into CAPI. Then, a certain number of households is selected from interviews completed per enumerator. For example, let's say in a week, an enumerator A completed 50 interviews, 10-15% of that – 5 or 7 households from the lists already visited will be revisited by a data validator (M&E staff of SOCU). This is done for every enumerators and the households are selected randomly. This provides a feedback loop for data administrators on how efficient an enumerator is.	
8	Backend Analysis	Data	The backcheck data analysis requires matching initial data collected with backcheck data on the same households. This will produce a report that will show level of discrepancy in data collection. The report is based on each data enumerators and will be used to assess the enumerators' activities on the field. Further, the backcheck report will indicate thresholds for commendation, warning, strikes and dismissal for enumerators.	
9	Backcheck and Validation	Support	This is a second layer support and report. The NASSCO MIS will provide support to the states on implementation of the backcheck process and the state will provide the report on backcheck to the federal as part of the M&E system on the data collection exercise.	

¹ A round of data collected could be within a week or two weeks. I will advise that data should not be left to be too large before data check begins. This will afford the team to make quick corrections whenever there is a need to go back to households.

10	Backend Cleaning	Data	<p>This requires the SOCU MIS team perform data consistency review and make corrections directly on the data. Likewise, the state MIS will perform all required data formatting, conversion and naming conventions that is required.</p> <p>At this stage, the PMT scores and its decile cut-offs will be generated and included as two fields in the data set. This is then passed to the federal team for validation.</p>	
11	Backend Validation	Data	<p>The federal team will validate the cleaned and formatted dataset from the states to be sure it conforms with required standards and rules prescribed by NASSCO. The NASSCO MISS will check consistence and correctness of the PMT and its decile cut-offs. The NASSCO MIS signs off on it and passes the data to the WB for clearance.</p>	
12	World Clearance	Bank	<p>The data is reviewed by the WB and the WB provides clearance for a limited time (first 6 months) after which the WB will cease to provide review and clearance but will do annual post audit of both SSRs and NSR.</p>	
13	State Social Registry		<p>The validated and cleared dataset will be stored aside as State Social Registry. Each dataset from each round of enumerations will be added up or appended to continue to build the SSRs.</p>	
14	National Registry	Social	<p>Each state will transfer their SSRs to the federal (NASSCO) to build the National Social Registry (NSR).</p>	